Cyber bullying Detection on Social Networks Using Machine Learning Approaches

MOHAMMED LIJAS C

MES20MCA-2028

PRODUCT OWNER: MR VASUDEVAN T.V

TABLE OF CONTENTS

INTRODUCTION

- Cyberbullying can be defined as aggressive, intentional actions performed by an individual or a group of people via digital communication methods such as sending messages and posting comments against a victim.

- Different from traditional bullying that usually occurs at school during face-to-face communication, cyber bullying on social media can take place anywhere at any time.

- Cyberbullying frequently leads to serious mental and physical distress, particularly for women and children, and even sometimes force them to attempt suicide.

- The purpose of this research is to design and develop an effective technique to detect online abusive and bullying messages by merging natural language processing and machine learning. Two distinct features, namely Bag-of - Words and term frequency-inverse text frequency (TF-IDF), are used to analyse the accuracy level of four distinct machine learning algorithms.

- In this paper, we proposed a novel friend recommendation framework (FRF) basedon the behavior of users on particular SNS's. The proposed method is consisted of the following stages: measuring the frequency of the activities done by the users and updating the dataset according to the activities, applying FP- Growth algorithm to classify the user behavior with some criteria, then apply multilayer thresholding for friend recommendation.

- In propsed system, we suggest a cyberbullying detection model based on machine learning that can detect whether a text relates to cyberbullying or not.or performance analysis, we use two different feature vectors BoW and TF-IDF.

- The results indicate that TF-IDF feature provides better accuracy than BoW. Users behaviour could be defined inseveral approaches like association rules in perspective of mining, complex graph activities, sequence mining etc.Suppose for two different user we have same behaviour we can recommend them each other.

MODULES

ADMIN

1. View user
2. Add bullying words
3. Add good words
4. View bullying words
5. View good words
6. View report

USER
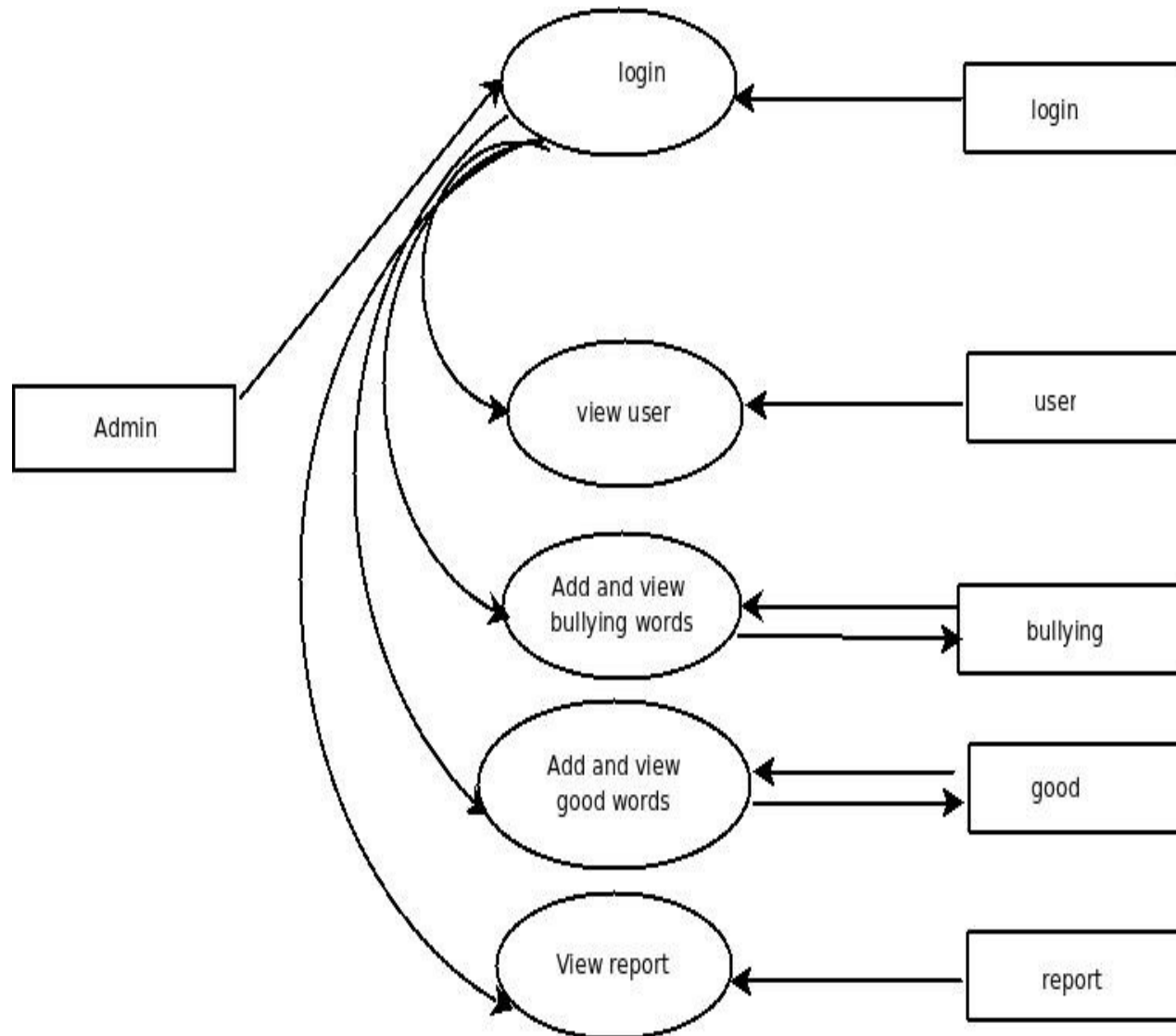
1. Registration

2. Add post

3. View my post

4. Chat

5. Add bullying words

6. Add friend request

7. View friend request

8. View recommendation

DATA FLOW DIAGRAM

LEVEL 0

LEVEL 1.1

Registration

login

add post

view post

chat

add and view
friend request

view recommendation

user

user

login

post

chat

request

comment
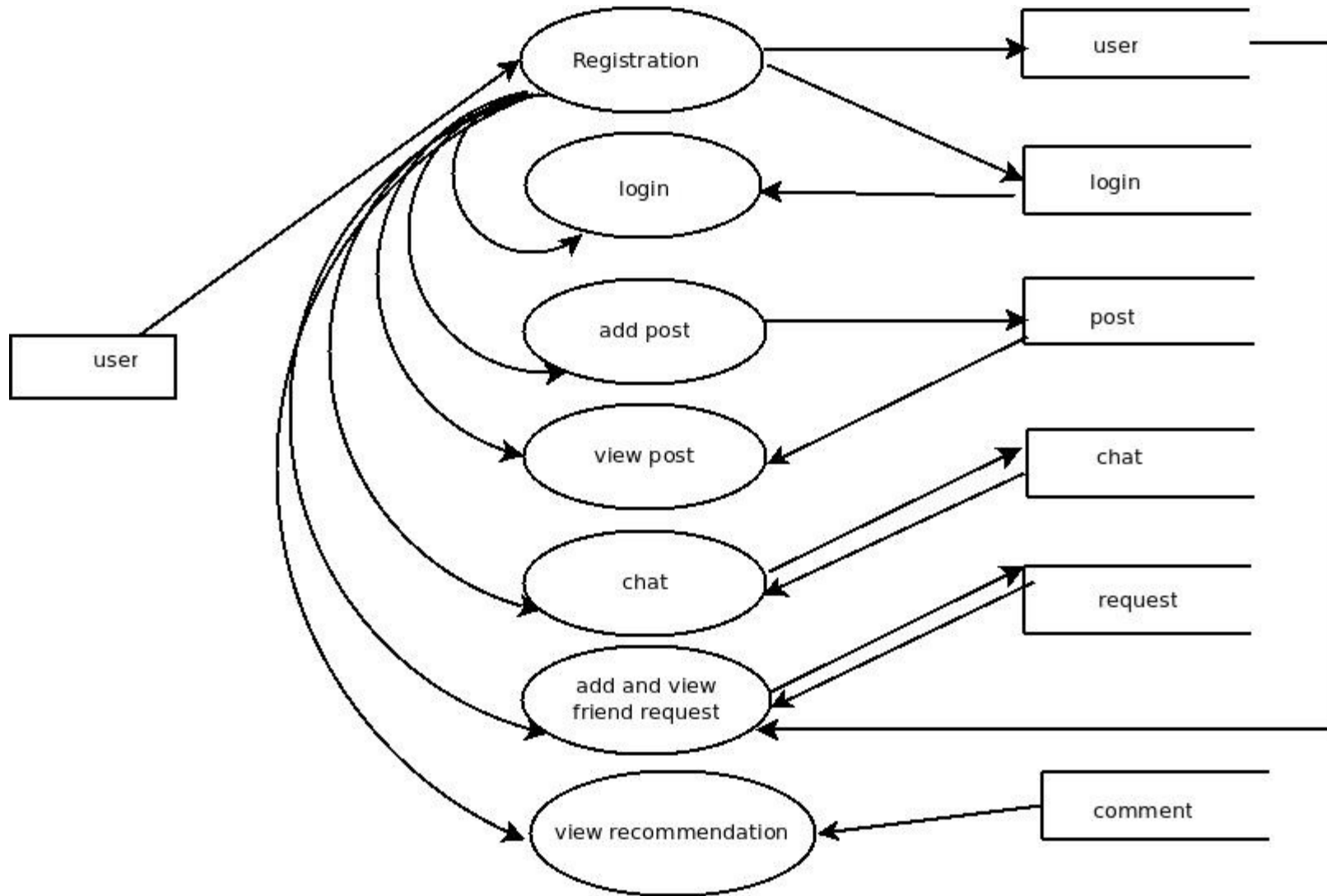
DEVELOPING ENVIRONMENT

Hardware Requirements

- Processor - Intel x86
- Speed - 1.1 GHz
- RAM - 4 GB (min)
- Hard Disk - 50 GB

Software Requirements

- Operating System - Windows 7 or Above , Linux
- Frontend – HTML,CSS,JavaScript
- Backend – Python ,MySQL
- Platform used - PyCharm, MySQL workbench
- Web Browser - Google Chrome, Fire fox, Microsoft Edge
- Frame work - Flask

PRODUCT BACKLOG

User Story ID	Priority <High/Medium/Low>	Size (Hours)	Sprint <#>	Status <Planned/In progress/Completed>	Release Date	Release Goal
1	Medium	2	1	Completed	08/01/2022	Table design
2	High	3		Completed	08/01/2022	Form design
3	High	5		Completed	08/01/2022	Basic coding
4	High	5	2	Planned		Creation of data set
5	Medium	5		Planned		Prediction
6	High	5	3	Planned		Filtering
7	high	5		Planned		Machine learning
8	Medium	5	4	Planned		Testing data
9	High	5		Planned		Output generation

USER STORIES

UserStoryID	As a <type of user>	I want to	So that I can
1	Admin	login	login successful with correct username and password
2	Admin	View users	View all user details
3	Admin	Add bullying words	Add bullying words
4			
5			
6			
7			
8			

PROJECT PLAN

User Story ID	Task Name	Start Date	End Date	Days	Status
1	Sprint 1	26/12/2021	28/12/2021	2	Completed
2		29/12/2021	31/12/2021	3	Completed
3		03/1/2022	08/01/2022	5	Completed
4	Sprint 2	09/01/2022	16/01/2022	8	Planned
5		18/01/2022	22/01/2022	5	Planned
6	Sprint 3	23/01/2022	27/01/2022	5	Planned
7		30/01/2022	05/02/2022	7	Planned
8	Sprint 4	06/02/2022	10/01/2022	5	Planned
9		16/02/2022	19/02/2022	4	Planned

SPRINT BACKLOG PLAN

Backlog Item	Status & completion date	Original estimate in hours	Day1	Day2	Day3	Day4	Day5	Day6	Day7	Day8	Day9	Day10	Day11	Day12	Day13	Day14
User story #1,#2,#3,#4		hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs
Table design	28/12/2021	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Form design	31/12/2021	3	0	0	0	1	1	1	0	0	0	0	0	0	0	0
Basic Coding	08/01/2021	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
User story #5,#6,#7,#8																
Creation of data set	16/01/2022	5	1	1	0	1	0	1	0	1	0	0	0	0	0	0
Prediction	22/01/2022	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Filtering	27/01/2022	5	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Machine learning	05/02/2022	5	0	0	0	0	0	0	0	1	0	1	1	1	0	1
Testing data	10/01/2022	5	1	1	1	1	1	0	0	0	0	0	0	0	0	0
User story #9																
Output generation	20/02/2022	5	0	0	0	0	0	0	0	0	0	2	1	1	1	1
Total		40	4	4	2	4	3	2	0	2	0	5	4	4	3	4

SPRINT ACTUAL

Backlog Item	Status & completion date	Original estimate in hours	Day1	Day2	Day3	Day4	Day5	Day6	Day7	Day8	Day9	Day10	Day11	Day12	Day13	Day14
User story #1,#2,#3,#4		hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs
Table design	28/12/2021	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Form design	31/12/2021	2	0	0	0	1	1	1	0	0	0	0	0	0	0	0
Basic Coding	08/01/2021	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
User story #5,#6,#7,#8																
Creation of data set																
prediction																
Filtering																
Testing data																
User story #9																
Output generation																
Total		10	1	1	0	1	1	1	0	0	0	1	1	1	1	1

THANK YOU